



**QUEEN'S
UNIVERSITY
BELFAST**

Validation of the completeness and accuracy of the Northern Ireland Cancer Registry

Kearney, T. M., Donnelly, C., Kelly, J. M., O'Callaghan, E. P., Fox, C. R., & Gavin, A. T. (2015). Validation of the completeness and accuracy of the Northern Ireland Cancer Registry. *Cancer Epidemiology*, 39(3), 401-404. <https://doi.org/10.1016/j.canep.2015.02.005>

Published in:
Cancer Epidemiology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2015, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/> which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Validation of the completeness and accuracy of the Northern Ireland Cancer Registry

Validation of the Northern Ireland Cancer Registry

Kearney TM¹, Donnelly C¹, Kelly JM¹, O'Callaghan EP¹, Fox CR¹, Gavin AT^{1*}.

1. Northern Ireland Cancer Registry, Queens University Belfast, Mulhouse building,
Grosvenor Road, BT12 6DP, UK

*Corresponding Author: Dr Anna Gavin: Northern Ireland Cancer Registry, Queens University Belfast, Mulhouse building, Grosvenor Road, BT12 6DP, UK.

email: a.gavin@qub.ac.uk

Tel: +44 28 9063 2573

This Northern Ireland Cancer Registry is funded by the Public Health Agency for Northern Ireland.

Word count 2092.

ABSTRACT

Background: It has been suggested that inaccuracies in cancer registries are distorting UK survival statistics. This study compared the Northern Ireland Cancer Registry (NICR) database of living patients, with independent data held by Northern Ireland's General Practitioners (GPs) to compare and validate the recorded diagnoses and dates held by the registry.

Methods: All 387 GP practice managers were invited to participate. 100 practices (25.84%) responded. Comparisons were made for 17,102 patients, equivalent to 29.08% of the living patients (58,798) extracted from the NICR between 1993-2010.

Results: There were no significant differences ($p>0.05$) between the responding and non responding GP patient profiles for age, marital status or deprivation score, however the responding GPs included more female patients ($p=0.02$). NICR data accuracy was high, 0.08% of GP cancer patients ($n=15$) were not included in registry records and 0.02% ($n=2$) had a diagnosis date which varied more than 2 weeks from GP records (3 weeks and 5 months). The NICR had recorded 2 different tumour types and 3 different tumour statuses (benign v malignant) to the GPs.

Conclusion: This comparison demonstrates a high level of accuracy within the NICR and that the survival statistics based on this data can be relied upon.

Key words: Validation, Cancer Registry completeness, Survival, unique patient identifier.

INTRODUCTION

Cancer registries are a fundamental component of cancer control programmes [1], with collated information used in the development, provision and evaluation of cancer services as well as public health decision making. The primary aim of cancer registration is to collect data using a standardised method to provide timely, high quality data which is nationally and internationally comparable on all incident cancer diagnoses within a well-defined population over time [2].

The Northern Ireland Cancer Registry (NICR) holds data on all cancers diagnosed in the population since 1993. There are three main sources of data utilised by the NICR; (i) histopathology and cytopathology, (ii) the Patient Administration System (PAS) currently used by all hospitals in Northern Ireland, which provides demographic information on patients and cancer diagnoses (ICD-10 codes) [3] and (iii) death certification. Between the year 1993 and 2010 histopathology, cytopathology and PAS combined were used to diagnose 98.97% of cancer cases, Death Certificate Initiated registrations were 3.57% out of which only 1.02% of diagnoses in these years were Death Certificate Only (DCO) notifications. Diagnoses are verified and confirmed by skilled Tumour Verification Officers (TVOs) who regularly scrutinise hospital records and histopathology text reports, whilst adhering to standardised international procedures [4]. The TVOs also have access to electronic data sources including the Clinical Oncology Information System (COIS) which contains annotated clinical oncology notes and the Cancer Patient Pathway System (CaPPs) [5]; a system which manages multidisciplinary cancer meetings.

The completeness and accuracy of the data held within cancer registries has recently been questioned [6,7], although suggestions for how to assess the completeness of registry data go

back as far as 1980 [8] . Beral and Peto [6] have suggested that as cancer registration is not compulsory within the United Kingdom (UK), unlike Scandinavian countries, [9,10] UK datasets are “somewhat incomplete”. They have speculated that UK survival statistics are distorted downwards due to the number of unregistered survivors, who will only be included on the registry upon death. These DCO registrations however, are excluded from survival analysis. Survival rates for many cancers are lower in the UK, compared to other countries within the western world [10,11] and it has been hypothesised that this may be due to unregistered cancer survivors or recurrences being mistakenly recorded as initial diagnoses.

Under the Quality and Outcome Framework (QOF); General Medical Services Contract (2004) [12], GPs are financially incentivised to maintain a high quality up-to-date record of patients with chronic conditions, including cancer. As well as receiving discharge summaries, outpatient letters and pathology and imaging reports [7], GPs also have direct contact with patients and have access to the patients’ detailed medical history. Consequently, they are an excellent data repository to use for the comparison and validation of the NICR data. This study aimed to assess the completeness and accuracy of the NICR database for surviving patients registered between 1993 and 2010, with cancer patient records held by Northern Ireland’s General Practitioners (GPs).

METHODS

A list of living patients, who received a malignant cancer diagnosis C00-C97 (WHO. 1992) between 1993 and 2010, was extracted from the NICR database and catalogued onto individual GP practice lists. Individuals, for whom a GP practice could not be assigned, were excluded. The GP patient lists were then encrypted and password protected. GP practice managers were contacted via email asking would they be willing to take part in the validation study. Practices

that consented were then forwarded two separate emails, one containing the encrypted spreadsheet relevant to their practice and the other the password. Practices were asked to compare the NICR list with their own practice list and to highlight differences between the two lists with regard to tumour type and stage, tumour site, date of diagnosis, changes to name and address and anything else which they felt was noteworthy. A variation of up to 2 weeks on either side of the diagnosis date was considered to be accurate. GP practices were also invited to notify the NICR of any cancer patients on their practice list but not on the NICR list.

Statistical Analysis

Chi square tests were used to make comparisons between responding and non responding GP practice patient profiles for gender, marital status, deprivation scores and tumour type (C00-C97) which were grouped together by site. An independent t-test was carried out to compare mean age of the individuals in the two groups. The alpha level was established as $p \leq 0.05$; values are reported as mean \pm SD.

RESULTS

387 GP practice managers were contacted, out of which 100 practices (25.84%) responded. Comparisons were made for 17,102 patients, equivalent to 29.08% of the NICR cohort (58,798 patients). 5,373 patients were excluded as the NICR had no record of their GP. There were no significant differences ($p > 0.05$) between the responding and non-responding GP patients for age, marital status and deprivation score (refer to table 1). No significant difference ($p > 0.05$), between the two groups for tumour type was detected (not tabulated). There were more female cancer patients among the responding GPs ($p = 0.02$) than the non-responding GPs. There was no significant difference ($p > 0.05$) in the size of GP practices between the responders and non-responders.

Table 2 presents the differences that were documented by the GP practices when comparing their lists to the NICR list. Out of 17,102 individuals compared, there was a variation on 2 patient tumour types and 3 patient designations as benign or malignant. It was later confirmed that the GPs' data with regard to status were correct and in each case the registry was updated accordingly. Date of diagnosis differed for 2 records, one record had a variation of 3 weeks and the other of 5 months. There were 15 individuals on the GP lists who were not included on the NICR.

2 of the 15 individuals lived outside of Northern Ireland when they were diagnosed and therefore would not normally be registered. A further 7 were resident in border counties and may have been diagnosed in the Republic of Ireland. COIS held data for 2 individuals but they had not been registered with the NICR. The NICR database was subsequently updated. The remaining 4 GP cases could not be verified and follow-up with the GP practices was unsuccessful.

DISCUSSION

This study compared two independent sources for 17,102 cancer patients and found a registry completeness of 99.9%. The researchers are confident in the representativeness of the sample as differences ($p > 0.05$) between responders and non-responders were not significant, except for patient gender. Although the feedback received was comprehensive indicating a sufficient time was spent comparing the lists, it must be taken into consideration, that some responding GPs may not have provided thorough comparisons between their data and that of the NICR.

The 4 individuals on the GP lists but not on the NICR database were diagnosed in different years, with different cancers and at different sites. Consequently, this small number of unregistered cancers would not materially influence the survival statistics of any particular cancer or year within Northern Ireland.

There were only 2 differences found in the recorded date of diagnosis between the NICR sample and the GP responses (a 3 weeks difference for a basal cell carcinoma and a 5 months difference for an infiltrating duct and lobular carcinoma). As the NICR require histopathology to confirm a diagnosis and GPs do not, it is reasonable to allow for a fortnight's discrepancy on either side of the NICR recorded diagnosis date. It can take up to 2 weeks for confirmation of diagnosis via histopathology, whereas GPs quite often record a diagnosis date earlier than the confirmed diagnosis date. However, in this study, the GP date of diagnosis was after the date recorded by the NICR, assuming that the NICR date of diagnosis is correct, there may have been a delay with the clinician informing the GPs. Although not conclusive, the results of this study could be taken to indicate that the date of diagnosis held within the NICR is the date of first diagnosis and not a recurrence.

Berkel [13] also conducted a similar study in the Netherlands with 11% GP participation and found that 15.4% of tumours recorded by the GPs were not on the cancer registry. In this study, only 0.1% of tumours on the GP records were not recorded by the NICR. The NICR also performed a similar (unpublished) study in 2004, which had a higher response rate (57%). This higher response was possibly due to the introduction at that time, of the Quality of Outcome Framework for GPs [12] and so the study was a means for GPs to ensure completeness of their own records as well as vice versa. In 2004, the GPs identified 2% of the cohort as having a different surname than the NICR records, this compares to 0.1% identified in 2010. 3 patients

were incorrectly categorised in 2004 by the NICR due to the same names being in the same household, e.g. a father/son mix-up, however in 2010 only one such error occurred. A high percentage (15%) of addresses required updating in 2004, compared to 0.45% in the current study. The high percentage of addresses which were amended in 2004, prompted the introduction of a successful scheme linking the NICR with GP patient registration, in order to regularly update patient demographic details. This link allowed for the follow-up of patients and also aided in participation in research studies such as in International Cancer Benchmark Partnership [10]. A recent publication [14] has indicated that 53 of the European General Cancer Registries claim to have $\geq 95\%$ completeness (self-assessed). However, the authors state that registries that did not rely on quantitative measures, such as the method of this study, were probably overly optimistic.

The improvements in the data's accuracy between 2004 and 2010 may be to some extent due to the introduction of the NHS unique identifier at the end of 2008. The unique identifier facilitates automatic matching of datasets thereby reducing the likelihood of missed changes to surnames or addresses and reduces the mix-ups between patients such as a father/son. It also enables the cancer registry to assess demographics thereby further improving the data held within the registry.

This study highlights the accuracy of the incidence records within the NICR as there were only 15 individuals included on GP lists which were not on the NICR and out of these 15, only 4 individuals could not be traced on any records. It is possible however, that these 4 individuals will be detected by the NICR through future death registrations. The low proportion of DCI (3.57%) and low DCO (1.02%) is also an indication of high data completeness within the NICR.

The results from this study disagree with the hypothesis that survival rates in Northern Ireland are underestimated due to underreporting of cancer survivors [6]. It can be concluded that the NICR holds an accurate database of living patients which can be relied upon for international comparisons regarding incidence and survival. Similar evidence for valid survival estimates in the UK has been presented by Woods [15] and Coleman[16], using a different methodology.

ACKNOWLEDGEMENTS

The authors would like to thank the Northern Ireland General Practice managers and their staff for taking part in this study, without their help, this report would not have been possible. The Northern Ireland Cancer Registry is funded by the Public Health Agency for Northern Ireland and has been granted ethical approval for its databases by ORECNI (10/NIR02/53).

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest or competing financial interests

Table 1: A comparison of cancer patients from responding and non responding GPs

	Non-responding GPs	Responding GPs	P value
GP Practices in border areas	13	14	0.65
Total patients registered	58798	17102	
Total cases	65758	19072	0.02
Male (%)	44.2	42.8	
Female (%)	55.8	57.2	
Age (years)	67.89 ± 15.52	66.79 ± 15.52	0.95
Patient marital status (%)			0.19
<i>Married</i>	43.3	43.3	
<i>Single</i>	8.3	8.6	
<i>Couple</i>	0.8	0.9	
<i>Widowed</i>	5.4	4.9	
<i>Unknown</i>	42.2	42.2	
Patient deprivation Score (%)*			0.20
<i>Quintile 1</i>	22.1	21.3	
<i>Quintile 2</i>	19.6	19.6	
<i>Quintile 3</i>	19.6	20.1	
<i>Quintile 4</i>	20.1	19.9	
<i>Quintile 5</i>	18	18.4	
<i>Unknown</i>	0.7	0.8	

* *Quintile 1 is the least deprived and quintile 5 is the most deprived.*

Table 2: Differences detected between the GP register of cancer patients and NICR database (n:17102)

Differences	Number of individual s	% of cohort (n17102)
Change of address	77	0.45
Different surnames	17	0.10
Incorrectly identified due to same name in household (Father/son)	1	0.005
Changed GP practice	168	0.98
GP different date of diagnosis	2	0.01
GP different tumour on record	2	0.01
GP different tumour status (benign/malignant)	3	0.02
GP no record of tumour	17	0.10
GP recorded death before 31/12/2010 (NICR unaware)	16	0.09
Not included on GP list*	110	0.64
<i>Not on NICR records</i>	15	0.08

*GPs indicated patients that were included on their list but not on the NICR list.

The NICR had a record of these patients, however they did not have the patients'

GP contact details, therefore these individuals were excluded from the study.

REFERENCES

- [1] Parkin, DM. The evolution of the population-based cancer registry. *Nat Rev Cancer* 2006;6:603-612, doi:10.1038/nrc1948
- [2] Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. The role of Cancer Registries. International Agency for Research on Cancer. Lyon: France.1999, Pp385-403
- [3] WHO. The ICD-10 International statistical classification of diseases and related health properties. 10th revision Vol 3. Geneva, World Health Organisation; 1994
- [4]Tyczynski JE, Demaret E, Parkin DM. Standards and guidelines for cancer registration in Europe. IARC Technical Publication. Vol 40. International Agency for Research on Cancer. Lyon: France; 2003
- [5] Napolitano G, Ranaghan E, Middleton R, Fox C, Gavin A. Real-time multidisciplinary team meeting management systems. A comparison with paper records for staging completeness and waiting times. *Behav Inform Technol* 2011;30:473-477, doi:10.1080/0144929X.2011.582151
- [6] Beral V, Peto R. UK cancer survival statistics are misleading and make survival worse than it is. *Br Med J* 2010;341:c4112, doi: 10.1136/bmj.c4112.
- [7] Boggon R, vanStaa TP, Chapman M, Gallagher AM, Hammad TA, Richards MA. (2013) Cancer recording and mortality in the General Practice Research Database and linked cancer registries. *Pharmacoepiemiol Drug Saf* 2011;22:168-175, doi: 10.1002/pds.3374
- [8] Goldberg J, Gelfand HM, Levy PS. Registry evaluation methods: a review and a case study. *Epidemiol Rev* 1980;2:210-220
- [9] Wagner G. *History of Cancer registration*, Eds Jenson OM, Parkin R, MacLennan CS, Muir CS, Skeet RG. Cancer Registration Principles and Methods. Lyon, France, International Agency for Research on Cancer, 1991, pp 3-6
- [10] Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, *et al.* Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995—2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. *Lancet* 2011;377:127-138, doi:10.1016/S0140-6736(10)62231-3
- [11] De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D *et al.* Cancer survival in Europe 1999—2007 by country and age: results of EURO CARE-5 - a population-based study. *Lancet Oncol* 2014;15: 23-34, doi:10.1016/S1470-2045(13)70546-1
- [12] Investing in General Practice, The New General Medical Services Contract (2004) www.legislation.gov.uk/uksi/2004/291/pdfs/uksi_20040291_en.pdf (accessed February 2015)
- [13] Berkel J. General Practitioners and completeness of cancer registry. *J Epidemiol and Community Health* 1990;44:121-124

- [14] Zanetti R, Schmidtmann I, Sacchetto L, Binder-Foucard F, Bordini A, Coza D et al. Completeness and timeliness: Cancer registries could/should improve their performance. *Eur J Cancer* 2014 doi: 10.1016/j.ejca.2013.11.040
- [15] Woods LM, Coleman MP, Lawrence G, Rashbass J, Berrino F, Rachet B. Evidence against the proposition that "UK cancer survival statistics are misleading": simulation study with National Cancer Registry data. *Brit Med J*. 2011;9;342:d3399. doi: 10.1136/bmj.d3399
- [16] Coleman MP, Rachet B, Woods L, Berrino F, Butler J, Capocaccia R, Dickman P, Gavin A, Giorgi R, Hamilton W, Lambert P, Peake MD, Perme MP, Stare J, Vedstedt P. Rebuttal to editorial saying cancer survival statistics are misleading. *Brit Med J*. 2011; 5:342-349 , d4214. doi: 10.1136/bmj.d4214

